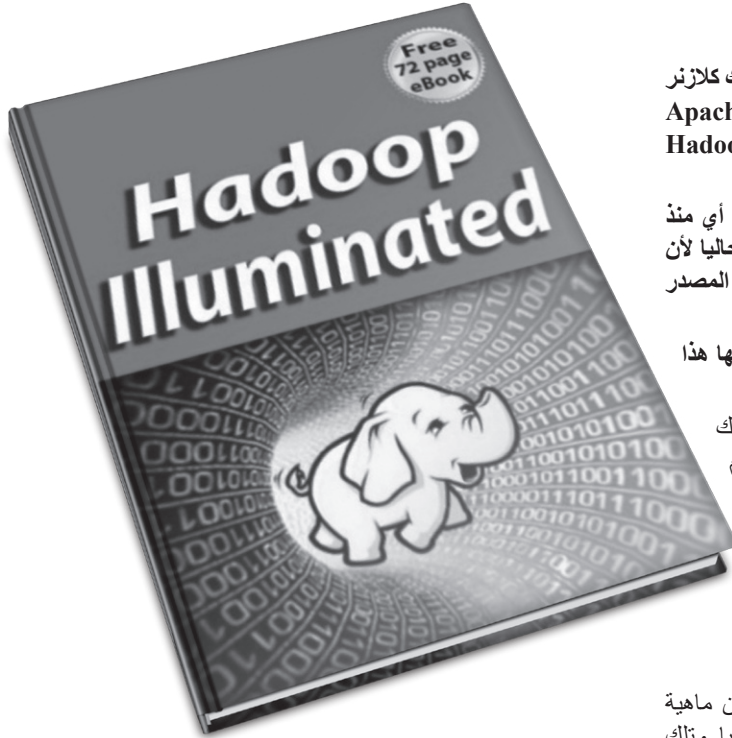


# «تكنولوجيا هادوب المضاءة أو المشعة»



تكنولوجيا هادوب المضاءة أو المشعة» تأليف كلا من مارك كلازير وسوجي ماتيا م، نشر عام ٢٠١٣ من قبل مؤسسة Apache Hadoop [https://github.cpm/Hadoop=illuminated-book]

وعلي الرغم من أن هذا الكتاب نشر أولا في عام ٢٠١٣ أي منذ خمسة سنوات مضت إلا أنه من المفيد التعرف علي أبعاده حاليا لأن تكنولوجيا الهادوب صارت تمثل إطار عمل برمجيات مفتوحة المصدر للتعامل مع ظاهرة البيانات الكبيرة الضخمة.

وقد تم استعراض خمسة عشر محورا أساسيا اشتمل عليها هذا الكتيب.

(١) لمن يقدم هذا الكتيب؟ تم استعراض للإجابة علي ذلك التساؤل استعراض أبعاد تكنولوجيا هادوب التي تتسم بالإشعاع الحالي.

(٢) لماذا يحتاج لإطار عمل هادوب؟ يحتاج لذلك لخسة أسباب رئيسية تتمثل في أن هادوب يقدم تخزين للبيانات الكبيرة الضخمة بتكلفة معقولة، كما يسمح لالتقاط بيانات جديدة وأكثر، ويمكنه تخزين البيانات مدة أطول، كما يقدم تحليلات ثرية الطابع.

(٣) البيانات الكبير الضخمة التمس استهلت بالتساؤل عن ماهية البيانات الكبيرة الضخمة، وتحديد البيانات المنتجة بشريا وتلك المنتجة آليا، مع توضيح من أين ترد البيانات الكبيرة، وبيان أمثلة بيانات كبيرة في العالم الواقعي، مع توضيح تحديات تلك البيانات الكبيرة الضخمة.

(٤) مقدمة رخوة Soft لتكنولوجيا هادوب، حيث استعرضت برامج MapReduce أو هادوب، ووضح لماذا أو بدائل هادوب، والبدايل الأخرى للحوسبة الموزعة الهائلة، مع بيان أبعاد الجدل الحادث عندئذ فيم يتعلق بتكنولوجيا هادوب.

(٥) بيان استخدام تكنولوجيا هادوب لرؤساء المنشآت والمطورين.

(٦) استعراض نظام ملف هادوب الموزع (HDFS) الذي تم تعريفه مفاهيمه وبيان معماريته.

(٧) مقدمة برامج MapReduce من خلال تحديث كيفية الفشل في تصميم المعالجة الموزعة، وبيان كيف أن برامج MapReduce يمكن عملها، وكيف يكون ذلك، وتحديد من اخترعها، والفوائد الناجمة من برمجتها.

(٨) استخدام حالات هادوب ودراسات حالات في المجالات التالية: السياسة، تخزين البيانات، الخدمات المالية، الرعاية الصحية، العلوم البشرية، الاتصالات عن بعد، الرحلات، الطاقة، واللوجسيات.

(٩) توزيع تكنولوجيا هادوب مع توضيح لماذا التوزيع، واستعراض مدي توزيع هادوب، واستخدام تكنولوجيا هادوب في الحوسبة السحابية.

(١٠) شرح نظام البيانات الكبيرة البيئي الذي يرتبط بالحصول

علي البيانات من نظام ملف هادوب الموزع (HDFS) وحساب إطار العمل المستخدم، مع تساؤل البيانات في نظام (HDFS) والتحميل في الحوسبة السحابية، وأدوات تدفق البيانات، ومراقبة النظم، والتصرفات، والتنسيق الموزع، وأدوات تحليلات البيانات، ومعالجة الرسالة الموزعة، مع أدوات ذكاء الأعمال، والمعالجة المتدفقة، وإطار عمل مبني علي YARN، مع متنوعات أخرى.

(١١) أدوات ذكاء الأعمال وارتباطها بتكنولوجيا الهادوب والبيانات الكبيرة الضخمة. حيث تم عرض حالة تلك الأدوات وتوضيح مصفوفة مقارنة خواصها وشرح للألفاظ المتضمنة.

(١٢) استعراض كلا من الأجهزة والبرمجيات التي ترتبط بمنصات تحميل تكنولوجيا الهادوب.

(١٣) بيان تحديات تكنولوجيا هادوب في نظام المنشأة البيئي، تحديد أن تكنولوجيا الهادوب ما زالت تتسم بالصعوبة حول حواف الأعمال.

(١٤) توافر بيانات كبيرة علانية من خلال مؤشرات لمجموعات البيانات ومستودعات بيانات جغرافية شاملة وبيانات الويب والبيانات الحكومية.

(١٥) الأخبار والوصلات مع البيانات الكبيرة الضخمة من خلال مواقع الأخبار والمدونات التي تتاح من قبل موردي وبناعي تكنولوجيا هادوب.